LEVEL II (12) B.S.

AD A058499

THE
GEORGE
WASHINGTON
UNIVERSITY

78 08 15 038

STUDENTS FACULTY STUDY R
ESEARCH DEVELOPMENT FUT
URE CAREER CREATIVITY CC
MMUNITY LEADERSHIP TECH
NOLOGY FRONTIE SIGN
ENGINEERING APP ENC
GEORGE WASHIN NIV

INSTITUTE FOR MANAGEMENT
SCIENCE AND ENGINEERING

SCHOOL OF ENGINEERING
AND APPLIED SCIENCE

LEVEL II

SOME CONTRIBUTIONS TO THE METHODOLOGY
FOR ANALYZING FAILURE DATA

by

Mahesh Chandra

Serial T-373
7 May 1978

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER T-373	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER

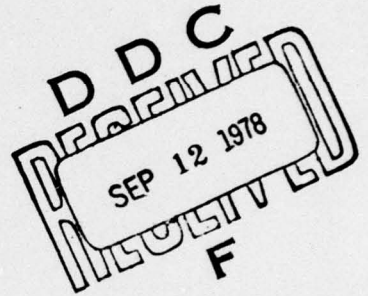| 4. TITLE (and Subtitle) SOME CONTRIBUTIONS TO THE METHODOLOGY FOR ANALYZING FAILURE DATA	5. TYPE OF REPORT & PERIOD COVERED SCIENTIFIC
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s) MAHESH CHANDRA	8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0263 NRC-04-78-239

9. PERFORMING ORGANIZATION NAME AND ADDRESS THE GEORGE WASHINGTON UNIVERSITY INSTITUTE FOR MANAGEMENT SCIENCE & ENGINEERING WASHINGTON, D. C. 20052	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH CODE 436 ARLINGTON, VIRGINIA 22217	12. REPORT DATE 7 May 1978
	13. NUMBER OF PAGES 68

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) NONE
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION OF THIS REPORT IS UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

SERIAL-T-373

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

FAILURE DATA ANALYSIS	GINI INDEX
GOODNESS OF FIT TESTS	TOTAL TIME ON TEST TRANSFORMS
WEIBULL DISTRIBUTION	MEAN RESIDUAL LIFE
ESTIMATED EMPIRICAL PROCESS	LORENZ PROCESS
LORENZ CURVE	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we deal with the important problem of "failure data analysis." Our contributions can be classified into two broad categories. Under the first category comes the development of some goodness-of-fit tests for the Weibull distribution with unknown parameters. Our test statistics are functions of the empirical distribution function, and are based on an adaptation of some recent results on the theory of goodness-of-fit tests. Under the second category comes our proposal that the (Cont'd)

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-014-6601 |

406 743

20. Abstract, Continued

"Lorenz curve" methods of economic theory also be considered for use in the analysis of failure data. In the sequel, we point out several interesting connectioons between some well-known indicators in economic theory and a central concept in reliability theory. We feel that pointing out these connections between the two apparently different disciplines constitutes an important contribution of this paper.

ACCESSION for

NTIS   White Section

DDC   Buff Section

UNANNOUNCED

JUSTIFICATION

BY

DISTRIBUTION/AVAILABILITY CODES

Dist.   SPECIAL

A

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

Abstract
of
Serial T-373
7 May 1978

SOME CONTRIBUTIONS TO THE METHODOLOGY
FOR ANALYZING FAILURE DATA

by

Mahesh Chandra

In this paper we deal with the important problem of "failure data analysis." Our contributions can be classified into two broad categories. Under the first category comes the development of some goodness-of-fit tests for the Weibull distribution with unknown parameters. Our test statistics are functions of the empirical distribution function, and are based on an adaptation of some recent results on the theory of goodness-of-fit tests. Under the second category comes our proposal that the "Lorenz curve" methods of economic theory also be considered for use in the analysis of failure data. In the sequel, we point out several interesting connections between some well-known indicators in economic theory and a central concept in reliability theory. We feel that pointing out these connections between the two apparently different disciplines constitutes an important contribution of this paper.

SOME CONTRIBUTIONS TO THE METHODOLOGY

FOR ANALYZING FAILURE DATA

By

Mahesh Chandra

B.Sc. 1963, Agra University, India

M.Sc. 1966, University of Delhi, India

A Dissertation submitted to

the Faculty of

The School of Engineering and Applied Science

of The George Washington University in partial satisfaction

of the requirements for the degree of Doctor of Science

May 7, 1978

Dissertation directed by

Nozer Darabsha Singpurwalla

Professor of Operations Research and

Research Professor of Statistics

Abstract

SOME CONTRIBUTIONS TO THE METHODOLOGY

FOR ANALYZING FAILURE DATA

by

Mahesh Chandra

Nozer Darabsha Singpurwalla, Director of Research

In this dissertation we deal with the important problem of "failure data analysis." Our contributions can be classified into two broad categories. Under the first category comes the development of some goodness-of-fit tests for the Weibull distribution with unknown parameters. Our test statistics are functions of the empirical distribution function, and are based on an adaptation of some recent results on the theory of goodness-of-fit tests. Under the second category comes our proposal that the "Lorenz curve" methods of economic theory also be considered for use in the analysis of failure data. In the sequel, we point out several interesting connections between some well-known indicators in economic theory and a central concept in reliability theory. We feel that pointing out these connections between the two apparently different disciplines constitutes an important contribution of this dissertation.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Chapter   Page

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION AND SUMMARY

In this dissertation we address the important problem of "failure data analysis." The significance of this problem for a practical implementation of reliability theory is too well known to warrant an explanation here.

A standard approach to the analysis of failure data is based on probability plotting methods, or the testing for "goodness of fit" [Mann, Schafer, and Singpurwalla (1974), pp. 214, 355]. Underlying the use of these methods is the assumption that the data constitute a *complete*, random sample from a fixed but unknown *parametric* family of distributions. Typical of these are the exponential, the Weibull, or the gamma distributions, and the problem is to test the hypothesis that the data have arisen from a specified member of the family.

Over the past few years significant interest has developed in failure data analysis, resulting in some imaginative approaches in the general area. For instance, Barlow and Campo (1975) have proposed the use of "total time on test plots" for a graphical analysis of failure data. The data could represent either a complete or an *incomplete* (censored or truncated) sample from a fixed but unknown *nonparametric* family of distributions. An example is the family of distributions whose failure rate is increasing [Barlow and Proschan (1975), p. 73]. Singpurwalla (1975) has proposed the use of time series techniques for analyzing data, which can be construed as the realization of a stochastic process. Such a model is appropriate when the failure data is correlated either due to contamination, to periodicities, or to the basic failure generating mechanism. The recent advances in testing for goodness of fit pertain to a

theory for tests based on the empirical distribution function when the parameters of the underlying failure distribution are estimated from the data. The pioneering work of Lilliefors (1967, 1969) sparked a flurry of papers around this general theme, the most recent ones contributed by Durbin (1973) and Serfling and Wood (1976). These latter two are unique in the sense that they emphasize a general theory for such tests, and are not directed towards a specific distribution.

In this dissertation we shall discuss our contributions to the general methodology for failure data analysis. Our contributions can be classified into two broad categories. Under the first we discuss the development of goodness-of-fit tests for the Weibull distribution with unknown parameters based on the empirical distribution function. Under the second category we propose that the "Lorenz curve" methods of economic theory be considered for use also in the analysis of failure data. In the sequel we point out several interesting connections between some well-known indicators in economic theory and a central concept in reliability theory. In agreement with several researchers in the two fields, we feel that pointing out the connections between the two apparently different disciplines constitutes an important and perhaps a major contribution of this dissertation.

A few words about the overall organization of this dissertation will be helpful to the reader. Chapters II and III are devoted to the problem of goodness-of-fit tests for the Weibull distribution with estimated parameters. In Chapter IV we discuss the Lorenz curve and the other measures of economic inequality; also discussed here are the relationships between these measures and a central concept in reliability theory. These relationships suggest the use of Lorenz curve methods for the analysis of failure data. Clearly, the theme of Chapters II and III is different from that of Chapter IV. Hence, Chapter IV can be perused independently of Chapters II and III. Some elements of the theory of the "weak convergence" of stochastic processes are relevant to both Chapters II and IV; these are presented in Chapter II. Literature relevant to the text of Chapters II and III is surveyed in Chapter II, whereas that which is relevant to the text of Chapter IV is surveyed in Chapter IV.

In what follows we summarize the major aspects of the material discussed in Chapters II, III, and IV.

In Chapter II we consider several test statistics based on the empirical distribution function, for testing the null hypothesis that a random sample belongs to a Weibull distribution with unknown scale and shape parameters. A foundation for testing such a hypothesis is provided by the fact that the logarithm of a Weibull random variable has an extreme value distribution with location and scale parameters, and by some recent results of Durbin (1973) and of Serfling and Wood (1976). These results pertain to the weak convergence of an associated "empirical" stochastic process under the null hypothesis. The asymptotic distribution of the empirical process serves as a basis for Monte Carlo studies to determine the appropriate critical points of the test statistics.

In Chapter III we give some results from a comparison of the power of our tests and an ad hoc but powerful test due to Mann, Scheuer, and Fertig (1973).

Chapter IV consists of several parts. We first show that the "Lorenz curve" and the "Gini index" are related to the "total time on test transform" and the "cumulative total time on test transform," respectively. Thus, the recently proposed tests for exponentiality based on the Gini statistic inherit the well-known properties of the tests for exponentiality based on the cumulative total time on test statistic.

Analogous to the "total time on test process" we define the "Lorenz process," and show its weak convergence to functionals of a Brownian motion process. This provides us with a theory for developing goodness-of-fit tests for any general distribution using the Lorenz curve and the Gini statistic. In addition to the above, we also state some new results on the geometry of the Lorenz curve that follow from the geometry of the total time on test transform.

In order to motivate the use of Lorenz curve methods for the analysis and interpretation of failure data, we show that there exists a

3

relationship between the "mean residual life" and the Lorenz curve. We illustrate our ideas by plotting and interpreting the Lorenz curves of two sets of failure data.

CHAPTER II

GOODNESS-OF-FIT TESTS FOR THE WEIBULL DISTRIBUTION

WITH ESTIMATED PARAMETERS

2.1 Introduction

The two-parameter Weibull distribution has found many applications in the engineering and in the biological sciences. For instance, it has been used by Cook, Doll, and Fellingham (1969) and by Doll (1971), to describe the observed age distribution of many human cancers. Its use for describing failures of electrical and mechanical components is well documented in the reliability literature.

In this chapter we address a fundamental problem involving any application of the Weibull distribution. We wish to test the null hypothesis that a given random sample belongs to a Weibull distribution with unknown parameters. Of the several methods for testing "goodness of fit," those based on the empirical distribution function are the most common. A foundation for these tests is the theory of weak convergence of stochastic processes. For the sake of completeness, we shall present in Section 2.2 the essential ingredients of this theory. In the sequel, we shall also introduce some notation and terminology.

2.2 Convergence of Stochastic Processes

In this chapter, as well as in Chapter IV, we will need to know the limiting behavior of certain stochastic processes that are of interest. In order to be able to do this satisfactorily, we shall have to introduce by way of preliminaries the following notations and definitions.

2.2.1 Preliminaries

Let C be the space of all continuous real functions on the closed unit interval $[0,1]$. We shall give C uniform topology by defining the distance between two functions x and y of $t \in [0,1]$ as

$$\rho(x,y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)| .$$

Let the class of Borel sets in C be denoted by $\mathscr{C}$. Then the *Wiener measure* W is a probability measure on $(C,\mathscr{C})$ with the following properties:

(i) For each $t \in [0,1]$, the random variable $x(t)$ is, under W, normally distributed with mean 0 and variance t; that is,

$$W\{x(t) \leq \alpha\} = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t} du .$$

If $t=0$, then $W\{x(0)=0\} = 1$.

(ii) For $0 \leq t_0 \leq t_1 \leq \cdots \leq t_k \leq 1$, the random variables

$$x(t_1)-x(t_0), \ x(t_2)-x(t_1), \ \ldots, \ x(t_k)-x(t_{k-1})$$

are independently distributed under W.

Billingsley (1968, p. 62) proves the existence of a Wiener measure on the space $(C,\mathscr{C})$.

We next consider some arbitrary probability space $(\Omega, \mathscr{B}, P)$, where $\mathscr{B}$ is the class of Borel sets in Ω, and P is a probability measure on $\mathscr{B}$. Let X be a P-measurable mapping from Ω into C; that is, $X^{-1}\mathscr{C} \subset \mathscr{B}$. Suppose that at any t, $t \in [0,1]$, the value of the mapping is denoted by $X(t,\omega)$, where $\omega \in \Omega$. Then, $\{X(t,\omega), 0 \leq t \leq 1\}$ is a stochastic process. It is called a *Wiener process* or a *Brownian motion process* if

$$P\{\omega: X(\omega) \in A\} = W(A), \quad A \in \mathscr{C} .$$

6

The mapping X from Ω into C is also known as a *random function* (if it can be measured).

A random function X in C is *Gaussian* if all its finite dimensional distributions are normal. The distribution of a Gaussian random function in C is completely specified by the means $E\{X(t,\omega)\}$, and the product moments $E\{X(t,\omega)X(s,\omega)\}$, $0 \leq s$, $t \leq 1$. Under a Wiener measure,

$$E\{X(t,\omega)\} = 0$$

and

$$E\{X(t,\omega)X(s,\omega)\} = s , \quad \text{if } s \leq t .$$

In order to study the behavior of empirical distribution functions, we shall need to define another random function of C , $Y(t,\omega)$, where

$$Y(t,\omega) = X(t,\omega) - tX(1,\omega) , \quad 0 \leq t \leq 1 .$$

Clearly Y is a Gaussian random function of C , and

$$E\{Y(t,\omega)\} = 0$$

and

$$E\{Y(t,\omega)Y(s,\omega)\} = s(1-t) , \quad \text{if } s \leq t .$$

The random function Y is called the *Brownian bridge*, or a *tied-down Brownian motion*. We also note that $Y(0,\omega) = Y(1,\omega) = 0$ with probability 1. The stochastic process $\{Y(t,\omega), 0 \leq t \leq 1\}$ is called the *Brownian bridge process*. The space C is not suitable to describe processes which contain jumps. We are thus led to consider a space which includes certain discontinuous functions.

The Skorokhod topology

Let D be the space of functions x on [0,1] that are right continuous and have left-hand limits:

(i) For $0 \leq t < 1$, $x(t^{+}) = \lim_{s \downarrow t} x(s)$ exists and $x(t^{+}) = x(t)$.

(ii) For $0 < t \leq 1$, $x(t^{-}) = \lim_{s \uparrow t} x(s)$ exists.

7

Let Λ denote the class of strictly increasing, continuous mappings of $[0,1]$ onto itself. If $\lambda \in \Lambda$, then $\lambda(0)=0$ and $\lambda(1)=1$.

For a pair of elements $x(t)$ and $y(t)$ of D, the *Skorokhod metric* $d(x,y)$ is defined to be the infimum of those positive ε for which there exists in Λ a λ such that

$$\sup_{0 \leq t \leq 1} |\lambda(t) - t| \leq \varepsilon$$

and

$$\sup_{0 \leq t \leq 1} |x(t) - y(\lambda(t))| \leq \varepsilon$$

[cf. Billingsley (1968, p. 111)].

We are interested in probability measures on $\mathscr{D}$, the Borel sets generated by the open sets of D. Billingsley (1968, p. 137) shows that the Wiener measure W which is defined on $(C, \mathscr{C})$ can be extended to $(D, \mathscr{D})$. Thus W can also be interpreted as a probability measure on $(D, \mathscr{D})$.

2.2.2 Convergence of Probability Measures

Consider arbitrary distribution functions F_n and F on the line. We say that F_n *converges weakly to* F, and denote this by $F_n \Rightarrow F$, if

$$F_n(x) \to F(x)$$

for all continuity points x of F. For example, if

$$F_n(x) = \begin{cases} 0, & \text{if } x < \frac{1}{n}, \\ 1, & \text{if } x \geq \frac{1}{n} \end{cases}$$

and

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases},$$

then $F_n \Rightarrow F$ even though $F_n(x) \not\to F(x)$ at $x = 0$.

8

The concept of weak convergence stated above is for the real line. It can also be formulated for a general metric space S . Suppose that P_n and P are probability measures on $\mathscr{S}$, where $\mathscr{S}$ is the class of the Borel subsets of S . Then, $P_n \Rightarrow P$, if and only if

$$\int_S f dP_n \to \int_S f dP , \quad \text{if } f \in C(S) ,$$

where $C(S)$ is the class of bounded, continuous, real-valued functions on S .

In order to discuss weak convergence in the space C we will have to consider what is known as the "tightness" of a sequence of probability measures $\{P_n\}$. The notion of tightness is too involved to present here, but it is explained in detail by Billingsley (1968, p. 54). For weak convergence in C , we state

<u>Theorem 2.2.1</u> [Billingsley (1968, p. 54)]: Let P_n and P be probability measures on $(C, \mathscr{C})$. If the finite dimensional distributions of P_n converge weakly to those of P , and if $\{P_n\}$ is tight, then $P_n \Rightarrow P$.

An analogous theorem by Billingsley (1968, p. 124) establishes conditions for the weak convergence of P_n to P in the space $(D, \mathscr{D})$.

Let $\{X_n\}$ be a sequence of random functions on C , and let X be a random function on C . We say that $\{X_n\}$ *converges in distribution to* X , written as

$$X_n \xrightarrow{\mathscr{D}} X ,$$

if the distributions P_n of X_n and P of X converge weakly; that is, if $P_n \Rightarrow P$.

9

2.2.3 Convergence of the Empirical
Distribution Function

Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be an ordered sample from a distribution F . The *sample*, or the *empirical distribution function (d.f.)* of F is defined as

$$F_n(u) = \begin{cases} 0 & , \quad u < X_{(1)} \\ i/n & , \quad X_{(i)} \leqq u < X_{(i+1)} \\ 1 & , \quad u \geqq X_{(n)} \end{cases} .$$

Let $F_n(t)$ denote the sample d.f. of a sample of n independent observations from the uniform distribution on $[0,1]$, $U[0,1]$. We define

$$Y_n(t) = \sqrt{n} \left(F_n(t) - t \right) , \quad 0 \leqq t \leqq 1 .$$

The stochastic process $\{Y_n(t), 0 \leqq t \leqq 1\}$ is called the *sample process* or *empirical process*. Note that $Y_n(t) \in D$. We can easily verify that

$$E[Y_n(t)] = 0$$

and

$$Cov[Y_n(s), Y_n(t)] = \min(s,t) - st , \quad 0 \leqq s , t \leqq 1 .$$

Thus the mean and the covariance of the process $\{Y_n(t)\}$ are identical to the mean and covariance of the Brownian bridge process $\{Y(t)\}$ discussed earlier. Furthermore, it can be shown [cf. Billingsley (1968, p. 141)] that the distributions P_n of Y_n converge weakly to the distribution P of Y . Thus, we may write

$$Y_n \xrightarrow{\mathscr{D}} Y . \tag{2.2.1}$$

A consequence of the above is that if g is a measurable function on D which is continuous almost everywhere with respect to the distribution

of $Y(t)$, then by the continuous mapping theorem of Billingsley (1968, Theorem 5.1), $g\big(Y_n(t)\big)$ converges in distribution to $G\big(Y(t)\big)$. As will be pointed out in the subsequent text, this result is useful for finding the asymptotic distributions of some test statistics used in testing for goodness of fit.

2.3 Goodness-of-Fit Tests Based on the Empirical Distribution Function for Testing Simple Hypotheses

Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be an ordered sample from a distribution $F(x)$. The goodness-of-fit test problem relates to the problem of testing the null hypothesis

$$H_0: \quad F(x) = F_0(x;\theta) \ ,$$

where θ is a vector of several parameters. The null hypothesis H_0 is called "simple" if $F_0(x;\theta) = F_0(x)$ is completely specified. We shall assume that $F_0(x)$ is continuous.

Let $F_0(x_{(i)}) = t_{(i)}$; then if the null hypothesis H_0 is true, $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(n)}$ is an ordered sample of size n from $U[0,1]$. Let $F_n(t)$ denote the empirical d.f. derived of $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$.

We shall consider test statistics based on the empirical d.f. We shall be concerned with two classes of test statistics. The first is typified by the Kolmogorov-Smirnov statistic

$$D_n = \sup_{0 \leq t \leq 1} \left| F_n(t) - t \right|$$

$$= \sup_{0 \leq t \leq 1} Y_n(t) \ ,$$

and the other is the Cramer-von Mises statistic

$$W_n^2 = n \int_0^1 [F_n(t) - t]^2 dt$$

$$= \int_0^1 Y_n^2(t) dt \ .$$

11

We observe that the above statistics can be written as functions of the empirical process $\{Y_n(t)\}$. Since $Y_n \xrightarrow{\mathcal{D}} Y$, where Y is the Brownian bridge process, $g\big(Y_n(t)\big) \xrightarrow{\mathcal{D}} g\big(Y(t)\big)$ if $g\big(Y(t)\big)$ is continuous in d for all $Y(t) \in D$. Thus the asymptotic distribution of $g\big(Y_n(t)\big)$ can be obtained from the asymptotic distribution of $g\big(Y(t)\big)$. For example, suppose that we are interested in obtaining the asymptotic distribution of $D_n = \sup\limits_{0 \leq t \leq 1} \big|Y_n(t)\big|$. It has been shown that $g\big(x(t)\big) = \sup\limits_{t} \big|x(t)\big|$ is continuous in d for all $x(t) \in D$, and that

$$P\left[\sup_{0 \leq t \leq 1} |Y(t)| \leq \alpha\right] = 1 - 2 \sum_{K=1}^{\infty} (-1)^{K+1} e^{-2K^2\alpha^2} , \quad \alpha \geq 0 ,$$

[cf. Billingsley (1968, p. 85)]. Thus

$$P\big[D_n \leq \alpha\big] \to 1 - 2 \sum_{K=1}^{\infty} (-1)^{K+1} e^{-2K^2\alpha^2} , \quad \alpha \geq 0 .$$

which is a classic result of Kolmogorov (1933).

By using similar arguments we can obtain the asymptotic distribution of W_n^2 .

There is another desirable feature of the statistic D_n . By the Glivenko-Cantelli theorem, when $F = F_0$,

$$P\left(\lim_{n\to\infty} D_n = 0\right) = 1 .$$

Thus, tests based on D_n are strongly consistent against all alternatives, i.e., as more observations are added, a false hypothesis is eventually rejected with probability one.

Other well-known goodness-of-fit statistics based on the empirical d.f. are:

the Kolmogorov-Smirnov one-sided statistics,

$$D_n^+ = \sqrt{n} \sup_{0 \le t \le 1} [F_n(t) - t]$$

$$= \sup_{0 \le t \le 1} [Y_n(t)] ,$$

$$D_n^- = \sqrt{n} \sup_{0 \le t \le 1} [t - F_n(t)]$$

$$= \sup_{0 \le t \le 1} [-Y_n(t)] ;$$

the Kuiper statistic,

$$D_n^{\pm} = D_n^+ + D_n^- ;$$

the Watson statistic,

$$U_n^2 = n \int_0^1 \left[F_n(t) - t - \int_0^1 \left(F_n(t) - t \right) dt \right]^2 dt$$

$$= \int_0^1 Y_n^2(t) dt - \left[\int_0^1 Y_n(t) dt \right]^2 ;$$

and the Anderson-Darling statistic,

$$A_n^2 = n \int_0^1 \frac{\left(F_n(t) - t \right)^2}{t(1-t)} dt$$

$$= \int_0^1 \frac{Y_n^2(t)}{t(1-t)} dt .$$

We also note that for a simple null hypothesis, the distributions of all the above statistics do not depend on $F_0(x)$; thus the tests based on these statistics are called "distribution-free tests."

13

2.4 Goodness-of-Fit Tests Based on the Empirical
Distribution Function with Estimated Parameters

We shall now consider the problem of testing the hypothesis

$$H_0: \quad F(x) = F_0(x; \theta) \; ,$$

where as before we assume that $F_0(x; \theta)$ is continuous; however, θ is a vector of unknown parameters which have to be estimated. For example, H_0 might be the hypothesis that data come from an exponential distribution with unknown mean. From a practical point of view, situations where the parameters are unknown are much more common than those where they are known.

Let $\hat{t}_{(j)} = F_0(x_{(j)}; \hat{\theta}_n)$, where $\hat{\theta}_n$ is a suitable estimator for the unknown parameter θ based on a sample of size n . Let $\hat{F}_n(t)$ denote the empirical distribution function derived from $\hat{t}_{(1)} \leq \hat{t}_{(2)} \leq \cdots \leq \hat{t}_{(n)}$. We are interested in those situations for which the distribution of $\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_n$ does not depend on θ . Otherwise, the significance points of the test statistics based on $\hat{F}_n(t)$ would depend on θ , which is unknown. If we restrict ourselves to the case where θ belongs to a location and scale family of distributions, and if we further assume that $\hat{\theta}_n$ is a complete sufficient statistic, then it follows from a theorem of Basu (1955) that $\hat{t}_{(1)}, \hat{t}_{(2)}, \ldots, \hat{t}_{(n)}$ have a distribution which is independent of θ .

Analogous to the test statistics for testing a simple hypothesis H_0 , we define the following modified test statistics:

(i) the modified Kolmogorov-Smirnov one-sided statistics,

$$\hat{D}_n^+ = \sqrt{n} \sup_{0 \leq t \leq 1} [\hat{F}_n(t) - t] \; ,$$

$$\hat{D}_n^- = \sqrt{n} \sup_{0 \leq t \leq 1} [t - \hat{F}_n(t)] \; ;$$

14

(ii)   the modified Kolmogorov-Smirnov statistic,

$$\hat{D}_n = \max\left(\hat{D}_n^+, \hat{D}_n^-\right) ;$$

(iii)   the modified Kuiper statistic,

$$\hat{D}^{\pm} = \hat{D}_n^+ + \hat{D}_n^- ;$$

(iv)   the modified Cramer-Von Mises statistic,

$$\hat{W}_n^2 = n \int_0^1 [\hat{F}_n(t) - t]^2 dt ;$$

(v)   the modified Watson statistic,

$$\hat{U}_n^2 = n \int_0^1 \left[\left(\hat{F}_n(t) - t\right) - \int_0^1 \left(\hat{F}_n(t) - t\right) dt\right]^2 dt ; \text{ and}$$

(vi)   the modified Anderson-Darling statistic,

$$\hat{A}_n^2 = n \int_0^1 \frac{\left(\hat{F}_n(t) - t\right)^2}{t(1-t)} dt .$$

Unlike the test statistics considered in Section 2.3, the modified test statistics presented above do not converge in distribution to functionals of a Brownian bridge process Y. However, their distribution does converge to a Gaussian process whose covariance depends on the assumed form for the null hypothesis $F_0(x;\theta)$ and the properties of $\hat{\theta}_n$. Thus the modified test statistics are not distribution-free, and do depend upon the form of $F_0(x;\theta)$. In the following section we discuss the asymptotic distribution of the modified test statistics when the assumed distribution is a Weibull.

2.5 Goodness-of-Fit Tests for the Extreme
Value Distribution Based on the
Estimated Empirical Process

The two-parameter Weibull distribution is given by

$$F(u; \delta, \beta) = 1 - \exp\left[-\left(\frac{u}{\delta}\right)^{\beta}\right], \quad u \geq 0$$

$$= 0 \qquad\qquad , \quad \text{otherwise,}$$

where the *scale* parameter δ and the *shape* parameter β are both assumed to be positive.

If we make the transformation $X = -\ln U$, where U has a two-parameter Weibull distribution, then the distribution of X is called the *extreme value distribution*. It is given by

$$G(x; a,b) = \exp\left\{-\exp\left[-\left(\frac{x-a}{b}\right)\right]\right\}, \quad b > 0 ,$$

where $a = -\ln\delta$ and $b = \frac{1}{\beta}$. We note that a and b are, respectively, the *location* and the *scale* parameters of the extreme value distribution.

To make a test of fit to the Weibull distribution, we shall first take the negative of the natural logarithms of the supposed Weibull data. Thus, we wish to test whether the distribution of a random sample $X_1, X_2,$ $\ldots, X_n$ is an extreme value distribution with unknown location parameter a and unknown scale parameter b. Specifically, we wish to test the "null hypothesis,"

$$H_0: \quad F(x) = G(x; \theta) = G(x; a,b) .$$

When a and b are known, then H_0 is simple and we can use the procedure discussed in Section 2.3. We shall now consider the case when a and b are unknown. Stephens (1977) has also considered tests for the extreme value distribution based on the empirical distribution function. However, his approach is completely different from ours and involves lengthy calculations utilizing numerical analysis. In Table 2.7.1 we shall show that our results compare favorably with his.

16

2.5.1 The Convergence Theorem and the
Modified Test Statistics

When a and b are not specified, that is, when H_0 is "composite," we consider an approach based on $(\hat{a}_n, \hat{b}_n)$, the maximum likelihood estimators of (a,b).

Let

$$\hat{G}_n(t) = \frac{1}{n} \sum_{i=1}^{n} I[G(X_i; \hat{a}_n, \hat{b}_n) \leq t] , \quad 0 \leq t \leq 1 ,$$

where $I[E]$ denotes the indicator of the event E. Analogous to the empirical process $\{Y_n(t)\}$ we define the estimated empirical process $\{\hat{Y}_n(t)\}$ as

$$\hat{Y}_n(t) = \sqrt{n} \, [\hat{G}_n(t) - t] , \quad 0 \leq t \leq 1 .$$

Our convergence theorem pertains to the estimated empirical process and is analogous to the result given by Equation (2.2.1). However, before stating the convergence theorem, we will have to introduce the following notation given in Durbin (1973), and verify that his conditions are satisfied.

Let us denote by θ the vector $[a,b]'$, and let θ_0 be any conveniently chosen value of θ. We state below a verification of the required conditions.

<u>Condition A</u>: The distribution $G(x, \theta_0)$ has a density $f(x, \theta_0)$ such that, for almost all x, the vector $\partial \ln f(x, \theta_0)/\partial \theta_0$ exists, and satisfies

$$E\left(\frac{\partial \ln f(x, \theta_0)}{\partial \theta_0} , \frac{\partial \ln f(x, \theta_0)}{\partial \theta_0'}\right) = \mathscr{J} ,$$

where $\mathscr{J}$ is finite and positive definite.

17

__Condition B__: Let $\hat{\theta}_n$ be the maximum likelihood estimator of θ; that is, $\hat{\theta}_n = [\hat{a}_n, \hat{b}_n]'$. Then, it is well known [cf. Cramer (1946)] that

$$n^{1/2}(\hat{\theta}_n - \theta_0) = \frac{1}{n^{1/2}} \mathscr{J}^{-1} \sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta_0)}{\partial \theta_0} + \varepsilon_n ,$$

where $\varepsilon_n \to 0$, in probability.

__Condition C__: Let N be the closure of a neighborhood of θ_0. Let $g(t, \theta) = \partial G(x, \theta)/\partial \theta$ when this is expressed as a function of t by means of the transformation $t = G(x; \theta)$; let $g(t) = g(t; \theta_0)$. The vector function $g(t, \theta)$ is continuous in (θ, t) for all $\theta \in N$, and $0 \leqq t \leqq 1$.

__Theorem 2.5.1__: By virtue of Conditions A, B, and C the *estimated empirical process* $\hat{Y}_n$ determined by the extreme value distribution $G(x; \hat{a}_n, \hat{b}_n)$ with $(\hat{a}_n, \hat{b}_n)$ as the maximum likelihood estimators, is such that

$$\hat{Y}_n \xrightarrow{\mathscr{D}} \hat{Y} ,$$

where $\hat{Y}$ is a Gaussian process in (D, d) with

$$E[\hat{Y}(t)] = 0 , \quad 0 \leq t \leq 1$$

and

$$E[\hat{Y}(s)\hat{Y}(t)] = \min(s,t) - st - g(s)' \mathscr{J}^{-1} g(t) , \quad 0 \leqq s , t \leq 1 .$$

$$(2.5.1)$$

__Proof__: The proof follows from Durbin (1973). //

If we choose $\theta_0 = [0,1]'$, then we have shown in Appendix A that $g(t) = [t\ln t, -t\ln t(\ln(-\ln t))]$; also

$$\mathscr{J}^{-1} = \begin{bmatrix} 1.10867 & 0.257 \\ 0.257 & 0.60793 \end{bmatrix},$$

[cf. Johnson and Kotz (1970, p. 282)]. Substituting the above into (2.5.1) we have the covariance of our Gaussian process

$$E[\hat{Y}(s)\hat{Y}(t)] = \min(s,t) - st - 1.108(s\ln s)(t\ln t)$$

$$+ 0.257(s\ln s)(t\ln t)\big(\ln(-\ln t)\big)$$

$$+ 0.257(s\ln s)\big(\ln(-\ln s)\big)(t\ln t) \qquad (2.5.2)$$

$$- 0.60793(s\ln s)\big(\ln(-\ln s)\big)(t\ln t)\big(\ln(-\ln t)\big), \quad 0 \leqq s, t \leqq 1.$$

Using the fact that if $h(\hat{Y}_n)$ is a function of $\hat{Y}_n$ which is continuous in metric d, $h(\hat{Y}_n) \overset{\mathcal{D}}{\longrightarrow} h(\hat{Y})$. Thus the limit laws of $\hat{D}_n^+$, $\hat{D}_n^-$, $\hat{D}_n$, $\hat{D}_n^{\pm}$, $\hat{W}_n^2$, $\hat{U}_n^2$, and $\hat{A}_n^2$ under H_0 are given, respectively, by the laws of the random variables

$$\hat{D}^+ = \sup_{0\leq t\leq 1} \hat{Y}(t)$$

$$\hat{D}^- = \sup_{0\leq t\leq 1} [-\hat{Y}(t)]$$

$$\hat{D} = \max(\hat{D}^+, \hat{D}^-)$$

$$\hat{D}^{\pm} = \hat{D}^+ + \hat{D}^-$$

$$\hat{W}^2 = \int_0^1 \big(\hat{Y}(t)\big)^2 dt \qquad (2.5.3)$$

$$\hat{U}^2 = \int_0^1 \big(\hat{Y}(t)\big)^2 dt - \left[\int_0^1 \hat{Y}(t)dt\right]^2,$$

and

$$\hat{A}^2 = \lim_{\varepsilon \to 0} \int_{0+\varepsilon}^{1-\varepsilon} \frac{\big(\hat{Y}(t)\big)^2}{t(1-t)} dt.$$

2.6 Asymptotic Distributions of the Modified Test Statistics

Monte Carlo methods were used to simulate the distribution of the limiting random variables given in (2.5.3). Following Serfling and Wood (1976) we approximate the Gaussian process by its finite-dimensional distributions, corresponding to an evaluation of the process at 29, 99, and 119

19

equally-spaced points in the unit interval. One thousand multivariate normal random vectors with the covariance given by Equation (2.5.2) were generated using a program from the International Mathematical and Statistical Library (IMSL). The empirical distributions of the supremum, the infimum, and the difference between the supremum and the infimum of the resulting multivariate normal vectors were then tabulated, thus approximating the limit laws of $\hat{D}_n^+$, $\hat{D}_n^-$, $\hat{D}_n$, and $\hat{D}_n^\pm$. Since the differences in the observed quantiles corresponding to the finite-dimensional distributions of $\hat{Y}$ at 29, 99, and 119 equally-spaced points diminished rapidly, the approximating procedure was terminated at 119 equally-spaced points. The asymptotic distributions of $\hat{W}_n^2$, $\hat{U}_n^2$, and $\hat{A}_n^2$ were obtained by using numerical integration techniques. For this we used Subroutine QSF from the IBM Scientific Subroutine Package. The various sample quantiles for the generated frequency distributions are shown in Table 2.6.1.

2.7 Concluding Remarks

As stated earlier, Stephens (1977) has also obtained the *asymptotic* percentage points for the statistics $\hat{W}_n^2$, $\hat{U}_n^2$, and $\hat{A}_n^2$. Stephens also gives a necessary modification so as to use these statistics for a finite sample size. Even though our approach is different, it is encouraging to note that our results seem to be in good agreement with those of Stephens. A comparison of the asymptotic points we obtained with those of Stephens is given in Table 2.7.1. Stephens has made no power comparisons, and since our results agree quite well with his, we conclude that our power comparisons given in Chapter III remain valid.

TABLE 2.6.1

ASYMPTOTIC DISTRIBUTIONS OF THE MODIFIED TEST STATISTICS
FOR EXTREME VALUE DISTRIBUTION WITH ESTIMATED PARAMETERS

p	pth Quantile						
	$\hat{D}_n^+$	$\hat{D}_n^-$	$\hat{D}_n$	$\hat{D}_n^\pm$	$\hat{W}_n^2$	$\hat{U}_n^2$	$\hat{A}_n^2$
0.010	0.241	0.261	0.320	0.595	0.016	0.016	0.106
0.025	0.274	0.278	0.343	0.645	0.019	0.018	0.136
0.050	0.296	0.306	0.369	0.681	0.022	0.021	0.155
0.100	0.326	0.330	0.408	0.738	0.025	0.024	0.181
0.250	0.395	0.399	0.477	0.837	0.036	0.034	0.238
0.500	0.487	0.478	0.560	0.977	0.051	0.049	0.333
0.750	0.601	0.583	0.669	1.157	0.073	0.069	0.454
0.900	0.722	0.707	0.785	1.305	0.105	0.098	0.623
0.950	0.785	0.787	0.841	1.408	0.123	0.117	0.746
0.975	0.850	0.853	0.910	1.497	0.147	0.140	0.849
0.990	0.908	0.977	0.981	1.614	0.175	0.164	0.991

TABLE 2.7.1

COMPARISON OF UPPER TAIL PERCENTAGE POINTS OF $\hat{W}_n^2$, $\hat{U}_n^2$, AND $\hat{A}_n^2$ STATISTICS

Statistic	$\alpha = 0.75$		$\alpha = 0.90$		$\alpha = 0.95$		$\alpha = 0.975$		$\alpha = 0.99$	
	Table 2.6.1	Stephens	Table 2.6.1	Stephens	Table 2.6.1	Stephens	Table 2.6.1	Stephens	Table 2.6.1	Stephens
$\hat{W}_n^2$	0.073	0.073	0.105	0.102	0.123	0.124	0.147	0.146	0.175	0.175
$\hat{U}_n^2$	0.069	0.070	0.098	0.097	0.117	0.117	0.140	0.138	0.164	0.165
$\hat{A}_n^2$	0.457	0.474	0.623	0.637	0.746	0.757	0.849	0.877	0.991	1.038

22

CHAPTER III

POWER COMPARISON FOR THE VARIOUS GOODNESS-OF-FIT

TESTS FOR THE WEIBULL DISTRIBUTION

3.1 Introduction

In order to evaluate the effectiveness of tests discussed in Chapter II, we evaluate their power against the lognormal distribution as an alternative. The lognormal distribution is chosen because it appears to be a natural competitor to a Weibull distribution. We also compare our tests with a test due to Mann, Scheuer, and Fertig (1973). We shall also consider some real failure data to illustrate the use of the modified test statistics.

3.2 Computing Formulas for the Modified Test Statistics

Let $\hat{\theta}_n$ be a suitable estimator of θ , and let $\hat{t}_{(i)} = F_0(X_{(i)}, \hat{\theta}_n)$, $i=1,2,\ldots,n$, where $X_{(1)} \leqq X_{(2)} \leqq \cdots \leqq X_{(n)}$ is the observed sample. We obtain the following formulas for computing the modified statistics discussed in Section 2.4.

23

$$\hat{D}_n^+ = \max_{1 \le i \le n} \left(\frac{i}{n} - \hat{t}_{(i)} \right)$$

$$\hat{D}_n^- = \max_{1 \le i \le n} \left(\hat{t}_{(i)} - \frac{i-1}{n} \right)$$

$$\hat{D}_n = \max(\hat{D}_n^+, \hat{D}_n^-)$$

$$\hat{D}_n^{\pm} = \hat{D}_n^+ + \hat{D}_n^-$$

$$\hat{W}_n^2 = \sum_{i=1}^n \left(\hat{t}_{(i)} - \frac{2i-1}{2n} \right)^2 + \frac{1}{12n}$$

$$\hat{U}_n^2 = \hat{W}_n^2 - n\left(\overline{t} - \frac{1}{2} \right)^2 \; ; \quad \overline{t} = \frac{1}{n} \sum_{i=1}^n \hat{t}_{(i)}$$

$$\hat{A}_n^2 = -\frac{1}{n} \sum_{i=1}^n (2i-1)\left(\ln \hat{t}_{(i)} + \ln\left(1 - \hat{t}_{(n+1-i)}\right) \right) - n \ .$$

$$(3.2.1)$$

For large samples, if the calculated value of a statistic in (3.2.1) exceeds the critical value at a given level of significance, then we reject the hypothesis H_0 at that level of significance.

For a finite sample of size n, the calculated values of the statistics obtained by using formulas (3.2.1) be multiplied by a factor of $(1 + 0.2/\sqrt{n})$ and then compared with the values given in Table 2.6.1. This factor was suggested by Stephens (1977) for $\hat{W}_n^2$, $\hat{U}_n^2$, and $\hat{A}_n^2$ and from our experience we have observed that this factor works well with the use of $\hat{D}_n^+$, $\hat{D}_n^-$, $\hat{D}_n$, and $\hat{D}_n^{\pm}$ for finite samples.

3.3 The Mann-Scheuer-Fertig (MSF) Test

The only other known procedure for testing goodness of fit for the Weibull that is not based on the empirical distribution function is a test proposed by Mann, Scheuer, and Fertig (1973).

24

The MSF test is based on a statistic S , and can be used for censored as well as uncensored samples. However, the percentage points of S and certain quantities that are used in calculating S are available only for sample sizes of up to 25. However, along with the necessary factor, the test statistics we have discussed can be used for any sample size.

For a sample of size n , censored at m , the statistic S is defined as

$$
S = \frac{\sum\limits_{i=[m/2]+1}^{m-1} \left(X_{(i+1)}-X_{(i)}\right) / [E(Y_{i+1}) - E(Y_i)]}{\sum\limits_{i=1}^{m-1} \left(X_{(i+1)}-X_{(i)}\right) / [E(Y_{i+1}) - E(Y_i)]} ,
$$

where $Y_i = \dfrac{X_{(i)}-a}{b}$ and $[r]$ denotes the greatest integer contained in r . Mann, Scheuer, and Fertig give percentage points of S and the values of the quantities $[E(Y_{i+1}) - E(Y_i)]$ for samples of size 3 to 25.

3.4 Power Calculations and Conclusions

The power comparisons were made numerically. For this random samples of size 20, 25, and 30, respectively, were generated from a lognormal (normal) distribution with parameters -0.5 (mean) and 1.00 (variance), respectively.

Maximum likelihood estimators of the parameters a and b of the extreme value distribution were obtained by numerically solving the following equations simultaneously:

$$
\hat{b} = \sum_j X_j/n - \left[\sum_j X_j \exp(-X_j/\hat{b}) \right] \left[\sum_j \exp(-X_j/\hat{b}) \right]^{-1} \qquad (3.4.1)
$$

and

$$
\hat{a} = -\hat{b}\ln\left[\sum_j \exp(-X_j/\hat{b})/n \right] . \qquad (3.4.2)
$$

The results of our power comparisons are shown in Tables 3.4.1, 3.4.2, and 3.4.3, and these are based on 1000 replicates. Based on this limited experiment, it appears that for samples of sizes 20 and 25, the MSF test has better power. For samples of size 30, the MSF test could not be used, and our modification of the Anderson-Darling test appears to have better power.

3.5 Example

Table 3.5.1 gives the failure data for the right rear brake on a D9G-66A Caterpillar tractor. Using total time on test plots (to be discussed in Chapter IV), Barlow and Campo (1975) conjecture that the data could have come from a Weibull distribution. We shall test this conjecture using our modified test statistics. Using the computing formulas (3.2.1), we obtain $\hat{D}_{107}^{+} = 0.505$, $\hat{D}_{107}^{-} = 0.453$, $\hat{D}_{107} - 0.505$, $\hat{D}_{107}^{\pm} = 0.957$, $\hat{W}_{107}^{2} = 0.050$, $\hat{U}_{107}^{2} = 0.049$, and $\hat{A}_{107}^{2} = 0.346$.

Comparing these values with those given in Table 2.6.1, we observe that they are not significant even at a level of significance of approximately 50%. Thus, based on our analysis, we confirm the conjecture of Barlow and Campo (1975) that the data could have arisen from a Weibull distribution.

TABLE 3.4.1

POWER COMPARISON: WEIBULL vs. LOGNORMAL
SAMPLE SIZE n=20

Level of Significance	Modified Kolmogorov-Smirnov			Modified Kuiper $\hat{D}^{\pm}_n$	Modified Cramer von Mises $\hat{W}^2_n$	Modified Watson $\hat{U}^2_n$	Modified Anderson Darling $\hat{A}^2_n$	Mann Scheuer Fertig S
	$\hat{D}^+_n$	$\hat{D}^-_n$	$\hat{D}_n$					
0.01	0.008	0.101	0.075	0.102	0.080	0.082	0.100	0.105
0.025	0.023	0.167	0.114	0.131	0.144	0.140	0.177	Not Available
0.05	0.046	0.236	0.171	0.209	0.219	0.211	0.238	0.265
0.10	0.084	0.418	0.249	0.303	0.322	0.321	0.354	0.432

TABLE 3.4.2

POWER COMPARISON: WEIBULL vs. LOGNORMAL
SAMPLE SIZE n=25

Level of Significance	Modified Kolmogorov-Smirnov			Modified Kuiper	Modified Cramer von Mises	Modified Watson	Modified Anderson Darling	Mann Scheuer Fertig
	$\hat{D}_n^+$	$\hat{D}_n^-$	$\hat{D}_n$	$\hat{D}_n^{\pm}$	$\hat{W}_n^2$	$\hat{U}_n^2$	$\hat{A}_n^2$	S
0.01	0.019	0.145	0.121	0.141	0.131	0.132	0.166	0.153
0.025	0.034	0.196	0.155	0.189	0.196	0.190	0.240	Not Available
0.05	0.072	0.263	0.206	0.255	0.279	0.264	0.310	0.389
0.10	0.116	0.457	0.286	0.360	0.364	0.365	0.411	0.533

TABLE 3.4.3

POWER COMPARISON: WEIBULL vs. LOGNORMAL
SAMPLE SIZE n=30

| Level of Significance | Modified Kolmogorov-Smirnov | | | Modified Kuiper | Modified Cramer von Mises | Modified Watson | Modified Anderson Darling | Mann Scheuer Fertig |
	$\hat{D}_n^+$	$\hat{D}_n^-$	$\hat{D}_n$	$\hat{D}_n^\pm$	$\hat{W}_n^2$	$\hat{U}_n^2$	$\hat{A}_n^2$	S
0.01	0.021	0.173	0.137	0.161	0.151	0.145	0.183	Not Available
0.025	0.042	0.252	0.185	0.210	0.231	0.219	0.286	Not Available
0.05	0.082	0.331	0.256	0.296	0.333	0.319	0.369	Not Available
0.10	0.132	0.521	0.355	0.412	0.421	0.418	0.475	Not Available

29

TABLE 3.5.1

FAILURE DATA FOR RIGHT REAR BRAKE
ON D9G-66A CATERPILLAR TRACTOR

56	1253	2325
83	1313	2337
104	1329	2351
116	1347	2437
244	1454	2454
305	1464	2546
429	1490	2565
452	1491	2584
453	1532	2624
503	1549	2675
552	1568	2701
614	1574	2755
661	1586	2877
673	1599	2879
683	1608	2922
685	1723	2986
753	1769	3092
763	1795	3160
806	1927	3185
834	1957	3191
838	2005	3439
862	2010	3617
897	2016	3685
904	2022	3756
981	2037	3826
1007	2065	3995
1008	2096	4007
1049	2139	4159
1069	2150	4300
1107	2156	4487
1125	2160	5074
1141	2190	5579
1153	2210	5623
1154	2220	6869
1193	2248	7739
1201	2285	

CHAPTER IV

THE USE OF THE LORENZ CURVE AND THE

GINI INDEX IN FAILURE DATA ANALYSIS

4.1 Introduction

A unifying concept in the statistical theory of reliability and life testing is the "total time on test transform," first discussed by Marshall and Proschan in 1965. Barlow (1968) and Barlow and Doksum (1972) have introduced and studied a scale-free test for exponentiality based on the "cumulative total time on test statistic," which is derived from the total time on test transform. Barlow and Campo (1975), and Barlow (1977) have studied the geometry of the total time on test transform, and have also used it for a graphical analysis of failure data. Langberg, Léon, and Proschan (1978) provide characterizations of the total time on test transform.

Measures of income inequality used by econometricians are the Lorenz curve and the Gini index (which is derived from the Lorenz curve). The Lorenz curve plots the percentage of total income earned by various portions of the population when the members are ordered by the size of their incomes. Gastwirth (1972) has studied the various properties of the Lorenz curve and the Gini index. Recently, in a series of papers, Gail and Gastwirth (1977a, 1977b) proposed scale-free tests for exponentiality based on the Lorenz curve and the Gini statistic. Among other things, they have shown that tests for exponentiality based on the Lorenz statistic and the Gini statistic are powerful against a variety of alternatives.

31

Our objective is to demonstrate that there exists a relationship between the total time on test transform and the Lorenz curve, and between the cumulative total time on test transform and the Gini index. Thus the tests for exponentiality proposed by Gail and Gastwirth (1977a, 1977b) inherit some very general properties of the tests for exponentiality based on the cumulative total time on test statistic given by Barlow and Doksum (1972). The relationship mentioned above also prompts us to define what we call the "Lorenz process" and discuss the weak convergence of this process to functionals of a Brownian motion process. Such a result increases interest in the Lorenz curve and the Gini index, since it provides a theory for developing goodness of fit tests for any general distribution using the Lorenz curve and the Gini index.

Gastwirth (1972) has given some properties of the geometry of the Lorenz curve that are of interest to an economist. In this chapter we present some additional properties of the geometry of the Lorenz curve, and generalize some of Gastwirth's results.

4.2 Definitions and Notation

Let X be a random variable with distribution F , and let μ be the mean of F ; let $F(0^-) \equiv 0$. Then, the *total time on test transform* is defined in

Definition 4.2.1:

$$H_F^{-1}(t) \overset{\text{def}}{=\!=} \int_0^{F^{-1}(t)} \overline{F}(u)\,du \,, \quad 0 \leq t \leq 1 \,,$$

where $\overline{F}(u) = 1-F(u)$ and $F^{-1}(t)$, the inverse of $F(t)$, is defined by

$$F^{-1}(t) = \inf_x \{x : F(x) \geq t\} \,.$$

It is easy to verify that $H_F^{-1}(1) = \mu$.

The *scaled total time on test transform* is defined in

Definition 4.2.2:

$$W_F(t) \quad \overset{\text{def}}{=} \quad \frac{H_F^{-1}(t)}{H_F^{-1}(1)} \ , \quad 0 \leq t \leq 1 \ .$$

In Figure 4.2.1, we show a plot of the scaled total time on test transform for a gamma distribution with shape parameters $\alpha = 1$ and 2 , respectively. Other properties of the scaled total time on test transform are discussed by Barlow and Campo (1975).

The *cumulative total time on test transform* is defined in

Definition 4.2.3:

$$V_F \quad \overset{\text{def}}{=} \quad \int_0^1 W_F(u)du = \frac{1}{\mu} \int_0^1 H_F^{-1}(u)du \ .$$

Thus, the cumulative total time on test transform is simply the area under the scaled total time on test transform.

Gastwirth (1971) has defined the *Lorenz curve* corresponding to a random variable X with distribution F , $F(0^-) = 0$, and mean μ as

Definition 4.2.4:

$$L_F(p) \quad \overset{\text{def}}{=} \quad \frac{1}{\mu} \int_0^p F^{-1}(u)du \ , \quad 0 \leq p \leq 1 \ .$$

In econometrics, $L_F(p)$ denotes the fraction of total income that the holders of the lowest pth fraction of incomes possess. In Figure 4.2.2, we show a plot of the Lorenz curve for a gamma distribution with shape parameters $\alpha = 1$ and 2 , respectively. It is easy to verify that the Lorenz curve is always a convex function of p .

Analogous to Definition 4.2.3, we define the *cumulative Lorenz curve* in

Definition 4.2.5:

$$(CL)_F \quad \overset{\text{def}}{=} \quad \int_0^1 L_F(p)dp = \frac{1}{\mu} \int_0^1 \int_0^p F^{-1}(u)du \ dp \ .$$
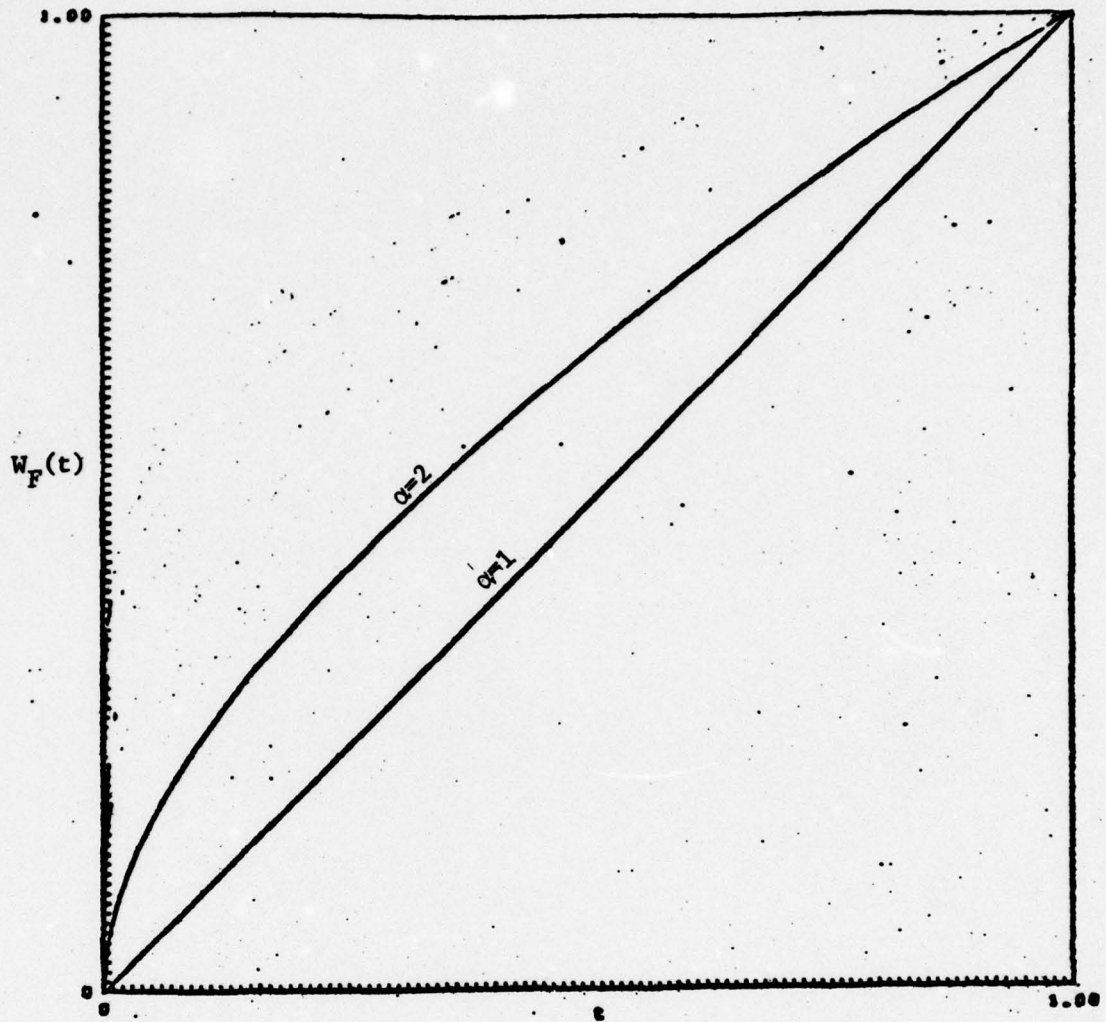
33

Figure 4.2.1--Total time on test transforms for gamma distribution
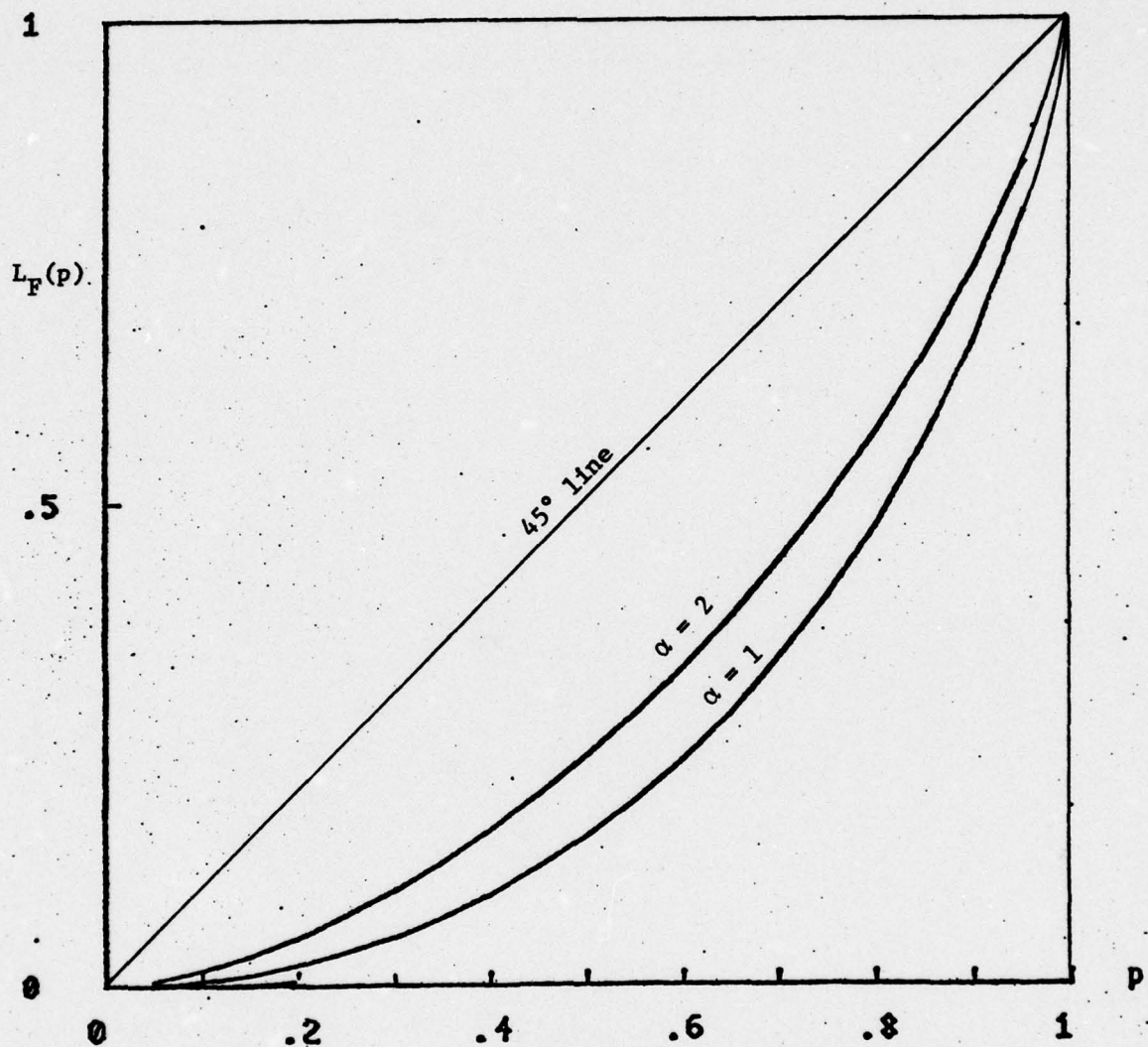
$$F(x) = \int_0^x \frac{\lambda^\alpha u^{\alpha-1} e^{-\lambda u}}{\Gamma \alpha} du \quad \text{for} \quad \alpha = 1,2 \ .$$

Figure 4.2.2--Lorenz curves for gamma distribution

$$F(x) = \int_0^x \frac{\lambda^\alpha u^{\alpha-1} e^{-\lambda u}}{\Gamma\alpha} \, du \quad \text{for} \quad \alpha = 1, 2 \, .$$

35

The most common measure of income inequality is the *Gini index*, defined in

Definition 4.2.6:

$$G_F = \frac{\frac{1}{2} - \int_0^1 L_F(p)dp}{(1/2)} \; .$$

That is, the Gini index is the ratio of the area between the Lorenz curve $L_F(p)$ and the 45° line, to the area under the 45° line (which is 1/2). The area between the 45° line and $L_F(p)$ is called the *area of concentration*.

4.3   Some Relationships Among the Concepts of Section 4.2

We now establish some relationships that exist among some of the concepts introduced in Section 4.2.

Integrating $W_F(t) = \frac{1}{\mu} \int_0^{F^{-1}(t)} \overline{F}(u)du$ by parts, we have

$$W_F(t) = \frac{1}{\mu}(1-t)F^{-1}(t) + \frac{1}{\mu}\int_0^t F^{-1}(u)du \; ,$$

or

$$W_F(t) = \frac{1}{\mu}(1-t)F^{-1}(t) + L_F(t) \; , \quad 0 \le t \le 1 \; . \qquad (4.3.1)$$

Thus, the scaled total time on test transform is related to the Lorenz curve as shown in (4.3.1).

Since $V_F = \int_0^1 W_F(t)dt$,

$$V_F = \frac{1}{\mu}\int_0^1 (1-t)F^{-1}(t)dt + \frac{1}{\mu}\int_0^1 \int_0^t F^{-1}(u)du \, dt \; .$$

Integrating by parts, we obtain

36

$$\int_0^1 \int_0^t F^{-1}(u)\,du\,dt = t \int_0^t F^{-1}(u)\,du \,\Big|_0^1 - \int_0^1 tF^{-1}(t)\,dt$$

$$= \int_0^1 F^{-1}(u)\,du - \int_0^1 tF^{-1}(t)\,dt$$

$$= \int_0^1 (1-t)F^{-1}(t)\,dt .$$

Thus

$$V_F = \frac{2}{\mu} \int_0^1 (1-t)F^{-1}(t)\,dt = \frac{2}{\mu} \int_0^1 \int_0^t F^{-1}(u)\,du\,dt ,$$

or

$$V_F = 2(CL)_F . \qquad (4.3.2)$$

Thus the cumulative total time on test transform is twice the cumulative Lorenz curve.

In order to see the relationship between the Gini index and the cumulative total time on test transform, we note that

$$G_F = 2\left[\frac{1}{2} - \int_0^1 L_F(p)\,dp \right]$$

$$= 1 - 2 \int_0^1 \frac{1}{\mu} \int_0^p F^{-1}(u)\,du\,dp ,$$

or

$$G_F = 1 - V_F . \qquad (4.3.3)$$

Thus the Gini index is simply one minus the cumulative total time on test transform.

Relationships (4.3.1), (4.3.2), and (4.3.3) now enable us to state some other results for the Lorenz curve and the Gini index.

4.4 Some Properties of the Lorenz
Curve and the Gini Index

Gastwirth (1972) has given several properties of the Lorenz curve and the Gini index that are of interest. We give here some additional properties which follow naturally from the results of the previous section.

Remark 4.4.1: L_F^{-1} , the inverse of L_F , is a distribution function with support on $[0,1]$; also L_F^{-1} is concave.

Proof: The conclusion follows from the fact that

$$L_F(1) = \frac{1}{\mu} \int_0^1 F^{-1}(p)dp = 1$$

when $F(0^-) \equiv 0$, and that $L_F^{-1}(p)$ increases in $p \in [0,1]$. Since L_F is convex, L_F^{-1} is concave.

We shall make use of Remark 4.4.1 in Theorem 4.4.6.

Definition 4.4.2: Let $\mathscr{F}$ be the class of *continuous distributions on* $[0,\infty)$, and let $\{deg\}$ be the class of degenerate distributions. Then F_1 *is star ordered with respect to* F_2 , denoted by $F_1 \underset{*}{\leqslant} F_2$, if $F_1, F_2 \in \mathscr{F} \cup \{deg\}$, and $\dfrac{F_2^{-1}F_1(x)}{x}$ is nondecreasing in x for $0 \leq x \leq F_1^{-1}(1)$.

We shall now state and prove

Theorem 4.4.3: If $F_1 \underset{*}{\leqslant} F_2$, and if $\int_0^\infty xdF_1(x) = \int_0^\infty xdF_2(x) = \mu$, then

(a) $L_{F_1}(p) \geq L_{F_2}(p)$

(b) $(CL)_{F_1} \geq (CL)_{F_2}$, and

(c) $G_{F_1} \leq G_{F_2}$.

38

Proof: Consider $L_{F_1}(p) - L_{F_2}(p) = \int_0^p \frac{1}{\mu}\left(F_1^{-1}(u) - F_2^{-1}(u)\right)du$; let

$h(u) = F_1^{-1}(u) - F_2^{-1}(u)$, and note that $\int_0^1 h(u)du = 0$. Since $F_1 \underset{*}{\leq} F_2$, by the "single crossing property" of star ordered distributions [cf. Barlow and Proschan (1975), p. 107], it follows that $h(u)$ changes sign exactly once, and from positive to negative values. Thus,

$$\frac{1}{\mu} \int_0^p h(u)du \geq 0 ,$$

and this completes part (a) of the theorem. Proof of parts (b) and (c) follow from the above result and the definitions of $(CL)_F$ and G_F .

If $F_1 \underset{*}{\leq} F_2$, and if F_2 is taken to be an exponential distribution, then F_1 belongs to the class of distributions which have "increasing failure rate average" [cf. Barlow and Proschan (1975)]. Theorem 7 of Gastwirth (1972) is analogous to Theorem 4.4.3 of this paper. However, our theorem is more general than that of Gastwirth, since it applies to a much larger class of distributions.

Definition 4.4.4: Let $\mathscr{F}$ be the class of continuous distributions on $[0,\infty)$, and let {deg} be the class of degenerate distributions. Then F_1 is *convex ordered with respect to* F_2 , denoted by $F_1 \underset{c}{\leq} F_2$, if $F_1, F_2 \in \mathscr{F} \cup \{deg\}$, and $F_2^{-1}F_1(x)$ is convex in x for $0 \leq x \leq F_1^{-1}(1)$.

Remark 4.4.5: $F_1 \underset{c}{\leq} F_2$ implies $F_1 \underset{*}{\leq} F_2$ [cf. Barlow and Proschan (1975, p. 107)].

In the following theorem we shall show that the convex ordering property is preserved by L_F^{-1} , the inverse of L_F . If $F_1 \underset{c}{\leq} F_2$, and

39

if F_2 is taken to be an exponential distribution, then F_1 belongs to the class of distributions which has an "increasing failure rate" [cf. Barlow and Proschan (1975)].

Theorem 4.4.6: If $F_1 \underset{c}{\leq} F_2$ then $L_{F_1}^{-1} \underset{c}{\leq} L_{F_2}^{-1}$.

Proof: We wish to show that $L_{F_2} L_{F_1}^{-1}(x)$ is convex in $0 \leq x \leq F_1^{-1}(1)$. We shall assume that F_1 and F_2 are absolutely continuous. Then we need only show that $\frac{d}{dx} L_{F_2}\left[L_{F_1}^{-1}(x)\right]$ is nondecreasing in $0 \leq x \leq F_1^{-1}(1)$.

Let μ_1 and μ_2 be the means of F_1 and F_2 , respectively. Then

$$\frac{d}{dx} L_{F_2}\left[L_{F_1}^{-1}(x)\right] = \frac{d}{dx} \frac{1}{\mu_2} \int_0^{L_{F_1}^{-1}(x)} F_2^{-1}(u)\,du$$

$$= \frac{F_2^{-1}\left(L_{F_1}^{-1}(x)\right)}{\mu_2} \frac{dL_{F_1}^{-1}(x)}{dx} .$$

Let

$$x = L_{F_1}(p)$$

$$\frac{dx}{dp} = \frac{F_1^{-1}(p)}{\mu_1}$$

$$\frac{dp}{dx} = \left. \frac{\mu_1}{F_1^{-1}(p)} \right|_{p = L_{F_1}^{-1}(x)} = \frac{\mu_1}{F_1^{-1}\left[L_{F_1}^{-1}(x)\right]} .$$

Hence

$$\frac{dL_{F_1}^{-1}(x)}{dx} = \frac{\mu_1}{F_1^{-1}\left[L_{F_1}^{-1}(x)\right]}$$

40

and

$$\frac{d}{dx} \frac{1}{\mu_2} \int_0^{L_{F_1}^{-1}(x)} F_2^{-1}(u)du \; = \; \frac{\mu_1 F_2^{-1}\left[L_{F_1}^{-1}(x)\right]}{\mu_2 F_1^{-1}\left[L_{F_1}^{-1}(x)\right]} \; .$$

Since $F_1 \underset{c}{\leq} F_2$ implies that $\dfrac{F_2^{-1}F_1(x)}{x}$ is nondecreasing in $0 \leq x \leq F_1^{-1}(1)$,

and since $F_1^{-1}L_{F_1}^{-1}(x) = t$ is nondecreasing in $0 \leq x \leq F_1^{-1}(1)$, a change of

variable shows that $\dfrac{d}{dx} \dfrac{1}{\mu_2} \displaystyle\int_0^{L_{F_1}^{-1}(x)} F_2^{-1}(u)du$ is nondecreasing in $0 \leq x \leq$

$F^{-1}(1)$. Since continuous distributions can be approximated arbitrarily
closely by absolutely continuous distributions, the proof of the theorem
is completed.

4.5 Some Statistics of Interest

Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics corre-
sponding to a random sample of size n from a distribution F , where
$F(0^-) = 0$. The *total time on test statistic* to the ith failure, $T\left(X_{(i)}\right)$,
is defined by

$$T\left(X_{(i)}\right) \; \overset{\text{def}}{=} \; \sum_{j=1}^{i} (n-j+1)\left(X_{(j)} - X_{(j-1)}\right) \; . \tag{4.5.1}$$

Barlow and Campo (1975) have used the *scaled total time on test*
statistic, $W\left(\dfrac{i}{n}\right)$, defined as

$$W\left(\frac{i}{n}\right) \; \overset{\text{def}}{=} \; \frac{\displaystyle\sum_{j=1}^{i} (n-j+1)\left(X_{(j)} - X_{(j-1)}\right)}{\displaystyle\sum_{j=1}^{n} X_{(j)}} \tag{4.5.2}$$

for analyzing failure data.

The *cumulative total time on test statistic*, V_n , defined as

$$V_n \overset{\text{def}}{=\!=} \frac{1}{n-1} \sum_{i=1}^{n-1} W\!\left(\frac{i}{n}\right) ,$$

has been used by Barlow and Doksum (1972) for testing for exponentiality. They show that a test based on V_n is asymptotically minimax against a class of alternatives defined by the Kolmogorov distance.

Gail and Gastwirth (1977a) have proposed a test for exponentiality based on the *Lorenz statistic*, $L_n(p)$, defined as

$$L_n(p) \overset{\text{def}}{=\!=} \sum_{i=1}^{[np]} X_{(i)} \Big/ \sum_{i=1}^{n} X_{(i)} ,$$

where $0 < p < 1$, and $[np]$ is the largest integer in np . The statistic $L_n(\cdot 5)$ is shown to have good power against a range of alternatives; this is based on a Monte Carlo investigation.

Recently, Gail and Gastwirth (1977b) proposed another test for exponentiality based on the *Gini statistic*, G_n , defined as

$$G_n \overset{\text{def}}{=\!=} \frac{\displaystyle\sum_{i=1}^{n-1} i(n-i)\left(X_{(i+1)} - X_{(i)}\right)}{(n-1)\displaystyle\sum_{i=1}^{n} X_{(i)}} . \tag{4.5.5}$$

Based upon Monte Carlo studies, Gail and Gastwirth (1977b) have concluded that G_n is *more powerful* than $L_n(\cdot 5)$ for $n=20$, against most of the alternatives that are studied.

We can easily verify the following relationships between the various test statistics that we have discussed thus far:

$$W\!\left(\frac{i}{n}\right) = L_n\!\left(\frac{i}{n}\right) + \frac{(n-i)X_{(i)}}{\displaystyle\sum_{j=1}^{n} X_{(j)}} \tag{4.5.6}$$

42

and

$$V_n = 1 - G_n . \qquad (4.5.7)$$

In view of (4.5.7) above, *the test for exponentiality based on the Gini statistic is identical to the test for exponentiality based on the cumulative total time on test statistic.* Thus, we can say that the test for exponentiality based on the Gini statistic is asymptotically minimax against some restricted alternatives.

The exact distribution of G_n under exponentiality follows from Theorem 6.2 of Barlow, Bartholomew, Bremner, and Brunk (1972), and from Equation (4.5.7) above. Gail and Gastwirth (1977b) have also derived the exact distribution of G_n, but by using a different argument.

4.6 The Lorenz Process and Its Weak Convergence

Using previous notation, we define the *Lorenz process*, $\left\{ \mathscr{L}_n(t) ; 0 \le t \le 1 \right\}$, as

$$\mathscr{L}_n(t) = \sqrt{n} \left\{ L_n\left(\frac{i}{n}\right) - L_F(t) \right\} , \qquad \frac{i-1}{n} < t \le \frac{i}{n}$$

$$1 \le i \le n ; \qquad (4.6.1)$$

$$\mathscr{L}_n(0) = 0 .$$

We are interested in the asymptotic behavior of this process for F in general.

4.6.1 Weak Convergence of the Lorenz Process

Let $\nu_n(u)$ be a discrete measure putting mass $1/n$ at $u = i/n$, $i = 1, 2, \ldots, n$. Then

$$L_n\left(\frac{i}{n}\right) = \frac{\frac{1}{n}\sum_{j=1}^{i} X_{(j)}}{\frac{1}{n}\sum_{j=1}^{n} X_{(j)}} = \int_0^{i/n} \frac{X_{([nu])}}{\overline{X}} \, d\nu_n(u) \, .$$

where $\overline{X} \stackrel{\text{def}}{=} \sum_{j=1}^{n} X_{(j)}/n$, and $[nu]$ is the greatest integer in nu .

Since $L_F(t) = \frac{1}{\mu}\int_0^t F^{-1}(u)du$, substitution in (4.6.1) gives

$$\mathscr{L}_n\left(\frac{i}{n}\right) = \int_0^{i/n} \sqrt{n}\left\{\frac{X_{([nu])}}{\overline{X}} - \frac{F^{-1}(u)}{\mu}\right\} d\nu_n(u)$$

$$+ \int_0^{i/n} \sqrt{n}\, \frac{F^{-1}(u)}{\mu} \, d[\nu_n(u) - u] \, .$$

(4.6.2)

If we assume that $\int_0^{\infty} x\, dF(x) < \infty$, and if $g = F^{-1}$ has a nonzero continuous derivative g' on $(0,1)$, then by Shorack (1972),

$$\sqrt{n}\left\{\frac{X_{([nu])}}{\overline{X}} - \frac{F^{-1}(t)}{\mu}\right\} \xrightarrow[\substack{n\to\infty \\ [nu]\to t}]{P} -\frac{g'(t)}{\mu} Y(t) - \frac{g(t)}{\mu^2} Z \, .$$

In the expression above "$\xrightarrow{P}$" denotes convergence in probability, Y is the Brownian bridge process on $(0,1)$, and $Z = \int_0^{\infty} Y[F(x)]dx$ is normal with mean 0 and variance σ_F^2 , where σ_F^2 is the variance of F .

Since the second term of (4.6.2) converges deterministically to zero, it follows that

$$\mathscr{L}_n(t) \xrightarrow[n\to\infty]{P} \int_0^t -\left(\frac{g'(u)}{\mu} Y(u) + \frac{g(u)}{\mu^2} Z\right) du \stackrel{\text{def}}{=} \mathscr{L}(t) \, .$$

We can also express $\mathscr{L}(t)$ as

$$\mathscr{L}(t) \;=\; -\,\frac{1}{\mu}\int_0^{F^{-1}(t)} Y[F(x)]dx \;-\; \frac{L_F(t)}{\mu}\int_0^{\infty} Y[F(x)]dx \;.$$

By a direct but tedious calculation [cf. Gail (1977)], it can be shown that under *exponentiality*

$$\mathrm{Var}\big(\mathscr{L}(t)\big) \;=\; 2(1-t)\ell n(1-t) + t + t(1-t) - \big(t + (1-t)\ell n(1-t)\big)^2 \;.$$

Thus, in contrast to an analogous result based on the convergence of the total time on test process [see Barlow and Campo (1975)], under exponentiality $\{\mathscr{L}(t)\,;\, 0 \leqq t \leqq 1\}$ is *not* the Brownian bridge.

4.6.2 *Uses of the Lorenz Process*

The Lorenz process can be used to find the asymptotic distribution of

$$\sup_{1 \leqq i \leqq n} \; \sqrt{n}\;\left| L_n\!\left(\frac{i}{n}\right) - L_F\!\left(\frac{i}{n}\right)\right| \;,$$

which by the invariance principle discussed in Section 2.2 is the same as that of

$$\sup_{0 \leqq t \leqq 1} \; \left| \mathscr{L}(t)\right| \;.$$

This statistic can be used to test the hypothesis that the given data has distribution F versus the general alternative that it does not. As seen at the end of Section 4.6.1, under exponentiality it is *not* the Kolmogorov-Smirnov statistic, and this is not very pleasing.

Another statistic that can be used for the same purpose is the area between the curve of $L_n\!\left(\frac{i}{n}\right)$ and the curve of $L_F(t)$. A consideration of this area leads us to Theorem 4.6.2, which follows from Theorem 6.6 of Barlow et al. (1972).

Theorem 4.6.1 [Barlow, Bartholomew, Bremner, and Brunk (1972)]: If

$$\int_0^\infty x\,dF(x) < \infty \quad \text{and} \quad \sigma^2(F) < \infty \,,$$

where

$$\sigma^2(F) = 2 \int\int_{s<t} \{2[1-F(s)] - V_F\} \times \{2[1-F(t)] - V_F\} F(s)\big(1-F(t)\big) \, ds \, dt \,,$$

then

$$\sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^{n-1} W\left(\frac{i}{n}\right) - V_F \right\} \xrightarrow[n\to\infty]{\mathcal{D}} N\left(0, \frac{\sigma^2(F)}{\mu^2}\right).$$

where "$\xrightarrow{\mathcal{D}}$" denotes convergence in distribution.

Using the fact that $G_F = 1-V_F$, and $G_n = 1-V_n$, we are now in a position to obtain the limiting distribution of the Gini statistic. We have

Theorem 4.6.2: Under the conditions of Theorem 4.6.1,

$$\sqrt{n} \left(G_n - G_F\right) \xrightarrow[n\to\infty]{\mathcal{D}} N\left(0, \frac{\sigma^2(F)}{\mu^2}\right).$$

In the case of F exponential, $G_F = \frac{1}{2}$ and $\sigma^2(F) = \frac{1}{12}$. Thus

$$\sqrt{12n} \left(G_n - \frac{1}{2}\right) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1) \,,$$

a result also obtained by Gail and Gastwirth (1977b) using some arguments due to Hoeffding (1949).

4.7 The Lorenz Curve and the Mean Residual Lifetime

Bryson and Siddiqui (1969) and also Hollander and Proschan (1975) have pointed out that the notion of "mean residual lifetime" is especially useful for the analysis of biological data. In this section we point out the relationship between the Lorenz curve and the mean residual lifetime.

46

Such a relationship suggests to us that the Lorenz curve methods, which have so far been mainly used in the social sciences, could also be used in the biological sciences. This possibility has also been hinted at by Thompson (1976).

The *mean residual lifetime* corresponding to a random variable X with distribution F, $F(0^-) \equiv 0$, is defined in

Definition 4.7.1:

$$\varepsilon_F(x) = \frac{\int_x^\infty \overline{F}(u)du}{\overline{F}(x)} .$$

We say that a distribution F has a *decreasing (increasing) mean residual lifetime* if $\varepsilon_F(x)$ is decreasing (increasing) in x for all $x \geq 0$.

Bryson and Siddiqui (1969) have used the decreasing mean residual lifetime property to interpret some survival data on patients suffering from leukemia.

If we denote the mean of F by μ, then we can write $\varepsilon_F(x)$ as

$$\varepsilon_F(x) = \frac{\mu\left[1 - \frac{1}{\mu}\int_0^x \overline{F}(u)du\right]}{\overline{F}(x)} .$$

From Definition 4.2.2, it follows that

$$\frac{1}{\mu}\int_0^x \overline{F}(u)du = W_F(F(x)) ;$$

thus

$$\varepsilon_F(x) = \frac{\mu\left[1 - W_F(F(x))\right]}{\overline{F}(x)} .$$

The above expression when used with Equation (4.3.1) gives us a relationship between $\varepsilon_F(x)$ and $L_F(\cdot)$; specifically, we have

47

$$L_F(F(x)) = 1 - \frac{\overline{F}(x)}{\mu} [\varepsilon_F(x) + x] . \qquad (4.7.1)$$

In order to demonstrate the use of the Lorenz curve for biological applications, we shall consider the data given by Bryson and Siddiqui (1969). These data pertain to survival times (in days), from the time of diagnosis, of patients suffering from chronic granulocytic leukemia. The ordered 43 survival times in days are: 7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495, 497, 532, 571, 579, 581, 650, 702, 715, 779, 881, 900, 930, 968, 1077, 1109, 1314, 1334, 1367, 1534, 1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509.

If we denote the number of survivors at time x by S, and if the size of the initial population is denoted by n, then Bryson and Siddiqui estimate the mean residual life at time x by

$$\hat{\varepsilon}(x) = S^{-1} \Sigma (X_j - x) ,$$

where X_j denotes the survival time of the jth element and the sum is for those having survived up to time x.

In Figure 4.7.1 we show a plot of $\hat{\varepsilon}(x)$ versus the time x, for the data in question. Thus, the distribution of survival times has a decreasing mean residual life; this conclusion is based upon an inspection of Figure 4.7.1.

In Figure 4.7.2 we give a plot of the sample Lorenz curve for these data. The sample Lorenz curve is simply a plot of the Lorenz statistic $L_n(p)$ (defined in Section 4.5) versus p, $0 < p < 1$. The sample Lorenz curve $L_n(p)$ represents the proportion of the total lifetime contributed by the least fortunate $p \cdot 100$ percent of the patients; for example, 50% of the patients contribute only 20% of the total lifetime. The sample Lorenz curve can also be used to compare the heterogeneity of the survival patterns of two groups of patients. To illustrate this, we give in Figure 4.7.3 the Lorenz curves for the data on the survival times of guinea pigs considered by Doksum (1974). The Lorenz curve for the "control
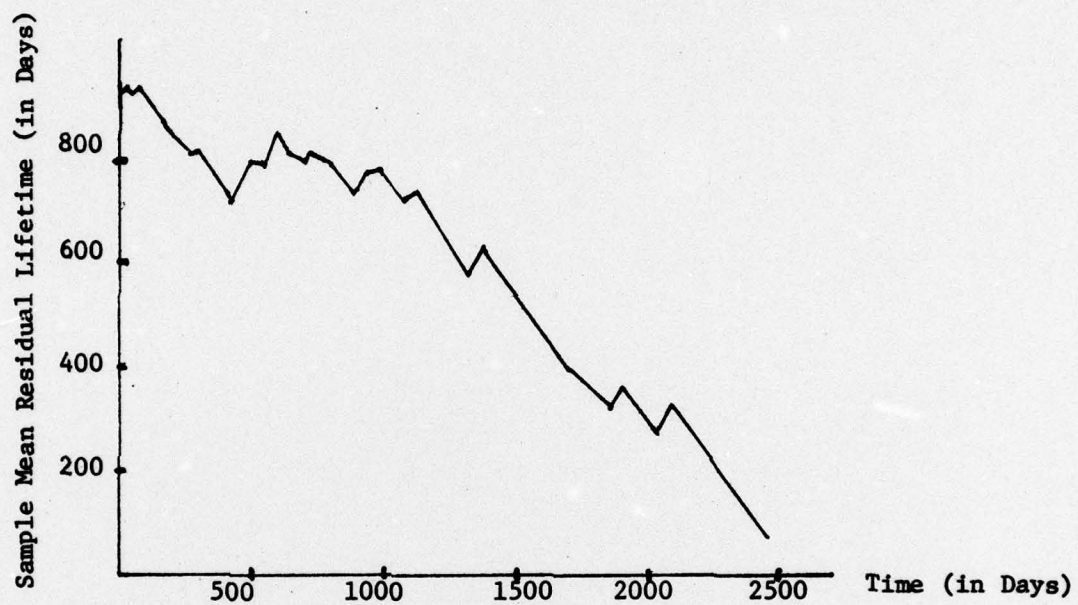
48

Figure 4.7.1--Sample mean residual lifetime versus
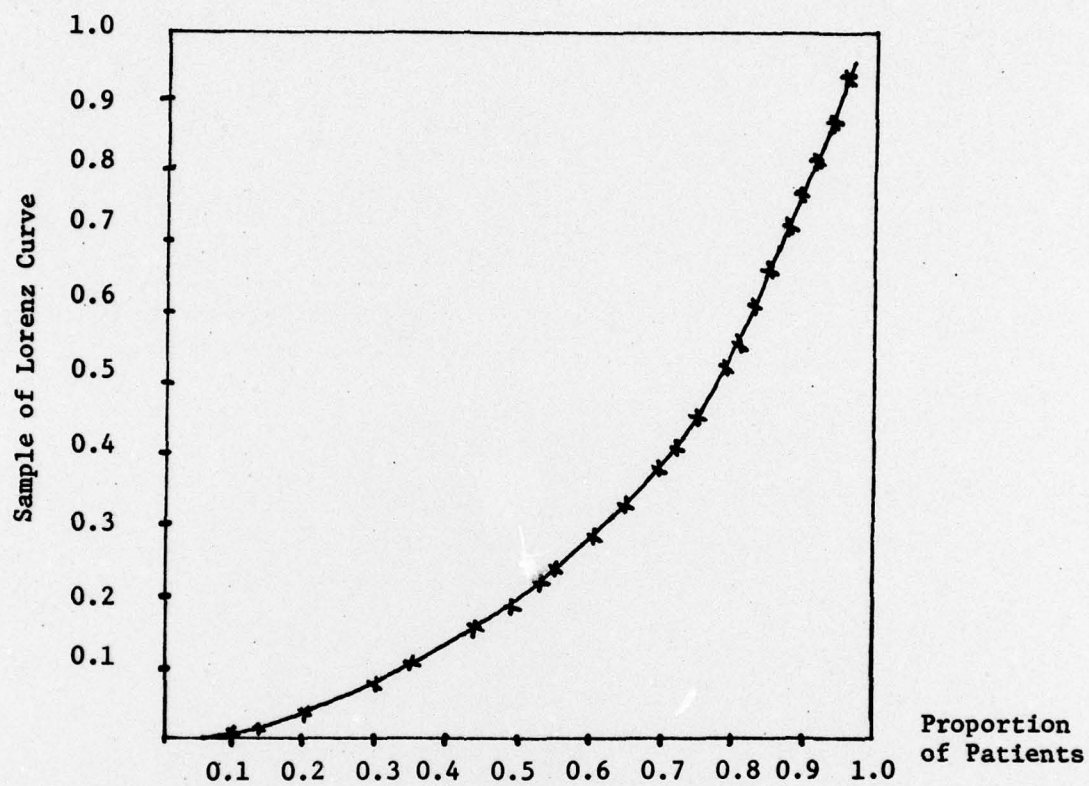time of leukemia patients

Figure 4.7.2--Sample Lorenz curve versus proportion
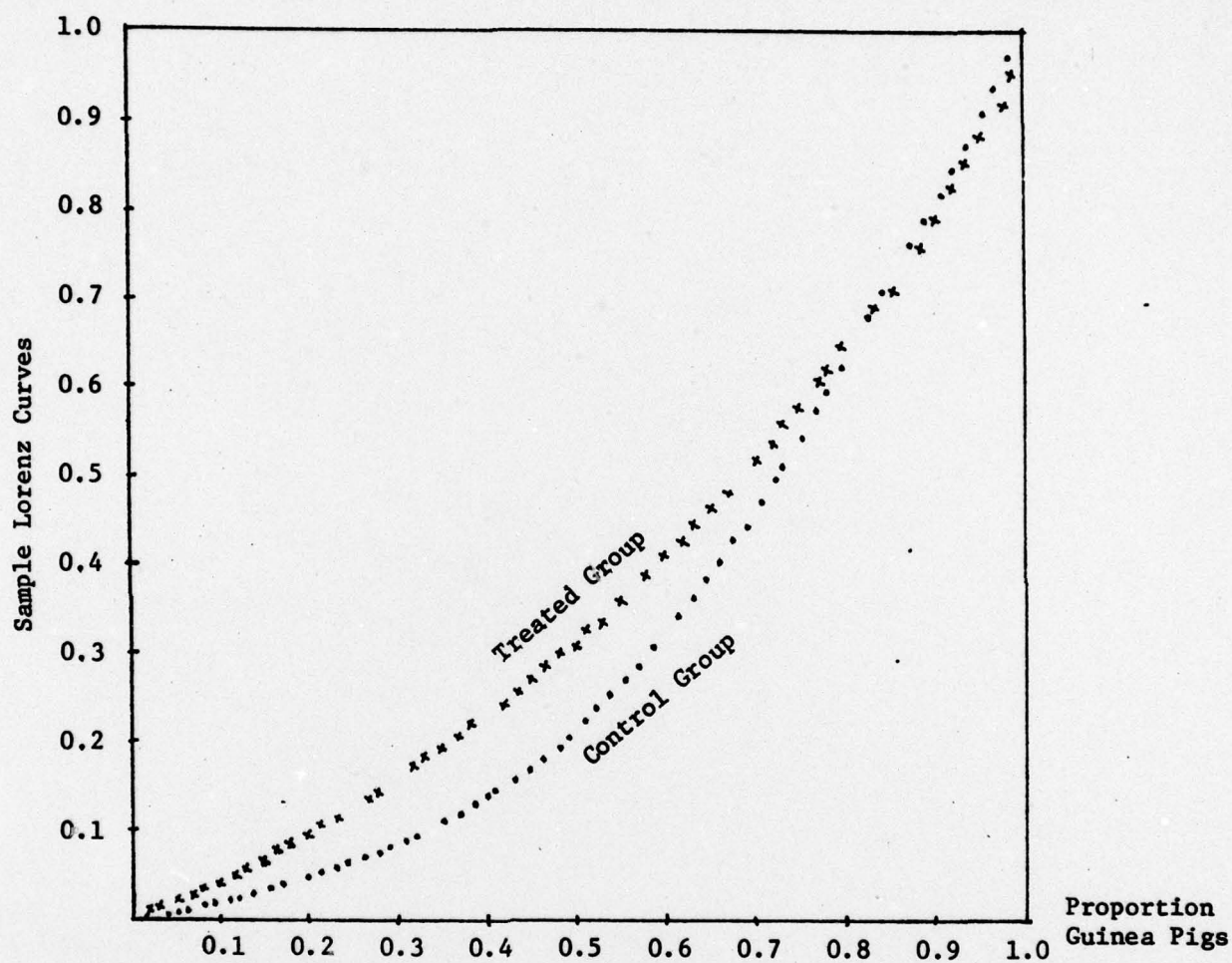of leukemia patients

Figure 4.7.3--Sample Lorenz curves versus proportion
of guinea pigs

group" lies below the Lorenz curve for the "treatment group" for p less
than about .8 ; the curves cross near $p \approx .8$. Thus, initially the
treatment group is less heterogeneous than the control group and the re-
verse is true later on. This can also be verified by an inspection of the
actual data.

CHAPTER V

RECOMMENDATIONS FOR FUTURE WORK

In Chapter II we have given the asymptotic percentage points of our modified test statistics for testing the goodness-of-fit statistics for the Weibull distribution with unknown parameters. Finding the exact percentage points of these test statistics for *finite samples* is an open and challenging question. Durbin (1975) is the only one who has given the exact percentage points for the modified statistics when the underlying distribution is an exponential. We find that his work involved a tremendous amount of theoretical and computational effort. We feel that finding the exact percentage points of the modified test statistics for the Weibull distribution would require a much greater effort. However, as we have stated in Chapter I, the analysis of failure data is an important problem, and thus such an effort would be well justified. This will also enable us to gain a better understanding of the power of our tests against various alternatives.

In Chapter IV, we have pointed out the connections among some well-known indicators in economic theory and a central concept in reliability theory. We hope this will help to consolidate and integrate statistical knowledge that has independently evolved in two different areas of application.

Up to this time the statistics based on the Lorenz curve, the Gini index, and the total times on test transforms have been used to develop tests for exponentiality. It is well known that the test for exponentiality based on these statistics have more power against most of the alternatives than the tests based on the empirical distribution function. We feel that tests for the gamma and the Weibull distributions based on

the Lorenz curve may prove to be more powerful than those proposed here.
Our results on the weak convergence of the Lorenz process has already given
us the necessary theory to pursue the development of such tests. Another
possible direction of research is the extension of the concepts of the
Lorenz curve and the Gini index to the multivariate case. Analogous
developments in reliability theory regarding the multivariate total time
on test transform could be used. This motivates us to propose a new con-
cept in economic theory, namely, a multivariate measure of social inequal-
ity. Thus there is a strong potential that further exploitation of the
connections that we have pointed out here will be possible.

APPENDIX A

DERIVATION OF $g(t)$

Let $G(x; a,b) = \exp\left\{-\exp\left[-\left(\frac{x-a}{b}\right)\right]\right\} = t$. Then

$$x = a - b \ln[-\ln t] .\qquad (A.1)$$

Let

$$g_a(x; a,b) = \frac{\partial G(x; a,b)}{\partial a}$$

$$= -\frac{1}{b} \exp\left\{-\exp\left[-\left(\frac{x-a}{b}\right)\right]\right\} \cdot \exp\left[-\left(\frac{x-a}{b}\right)\right]$$

and

$$g_b(x; a,b) = \frac{\partial G(x; a,b)}{\partial b}$$

$$= -\frac{(x-a)^2}{b^2} \exp\left\{-\exp\left[-\left(\frac{x-a}{b}\right)\right]\right\} \cdot \exp\left[-\left(\frac{x-a}{b}\right)\right] .$$

Then

$$g(x; a,b) = \begin{bmatrix} g_a(x; a,b) \\ g_b(x; a,b) \end{bmatrix} .\qquad (A.2)$$

From (A.1) and (A.2) we get

$$g(t) = g(t; 0,1) = \begin{bmatrix} t\ln t \\ -t\ln t(-\ln t) \end{bmatrix} .$$

55

REFERENCES

BARLOW, R. E. (1968). Likelihood ratio tests for restricted families of probability distributions. *Ann. Math. Statist.* 39 547-560.

BARLOW, R. E. (1977). Geometry of the total time on test transform. Operations Research Center Report ORC 77-11, University of California, Berkeley.

BARLOW, R. E., D. J. BARTHOLOMEW, J. M. BREMNER, and H. D. BRUNK (1972). *Statistical Inference under Order Restrictions.* John Wiley and Sons, Inc., New York.

BARLOW, R. E. and R. CAMPO (1975). Total time on test processes and applications to failure data analysis. *Reliability and Fault Tree Analysis* (R. E. Barlow, J. Fussell, and N. D. Singpurwalla, eds.). Conference volume published by SIAM, Philadelphia.

BARLOW, R. E. and K. DOKSUM (1972). Isotonic tests for convex ordering. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* 1 293-323.

BARLOW, R. E. and F. PROSCHAN (1969). A note on a test for monotone failure rate based on incomplete data. *Ann. Math. Statist.* 40 595-600.

BARLOW, R. E. and F. PROSCHAN (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart, and Winston, New York.

BASU, D. (1955). On statistics independent of complete sufficient statistics. *Sankhyā* 15 377-380.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley and Sons, Inc., New York.

BRYSON, M. C. and M. M. SIDDIQUI (1969). Some criteria for ageing. *J. Amer. Statist. Assoc.* 64 1472-1485.

COOK, P., R. DOLL, and S. A. FELLINGHAM (1969). A mathematical model for the age distribution of cancer in man. *Internat. J. Cancer* 4 93-112.

CRAMER, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, New Jersey.

DOKSUM, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Ann. Statist.* 2 267-277.

DOLL, R. (1971). The age distribution of cancer: implications for models of carcinogens. *J. Roy. Statist. Soc. Ser. A* 134 133-166.

DURBIN, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* 1 279-290.

DURBIN, J. (1975). Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings. *Biometrika* 62 5-22.

GAIL, M. H. (1977). A scale-free goodness-of-fit test for the exponential distribution based on the Lorenz curve. Ph.D. Dissertation, The George Washington University.

GAIL, M. H. and J. L. GASTWIRTH (1977a). A scale-free goodness-of-fit test for the exponential distribution based on the Lorenz curve. (Submitted for publication.)

GAIL, M. H. and J. L. GASTWIRTH (1977b). A scale-free goodness-of-fit

test for the exponential distribution based on the Gini statistic.

(Submitted for publication.)

GASTWIRTH, J. L. (1971). A general definition of the Lorenz curve.

Econometrica 39 1037-1039.

GASTWIRTH, J. L. (1972). The estimation of the Lorenz curve and Gini

index. *Rev. Economics and Statist.* 54 306-316.

HOEFFDING, W. (1949). A class of statistics with asymptotically normal

distribution. *Ann. Math. Statist.* 19 293-325.

HOLLANDER, M. and F. PROSCHAN (1975). Tests for the mean residual life.

Biometrika 62 585-593.

JOHNSON, N. L. and S. KOTZ (1970). *Continuous Univariate Distributions,*

1, Houghton-Mifflin Company, Boston.

KOLMOGOROV, A. (1933). Sulla determinazione empirica di una legge di

distribuzione. *Giorn. Ist. Ital. Attuari* 4 83-91.

LANGBERG, N., R. LEON, and F. PROSCHAN (1978). To appear.

LILLIEFORS, H. W. (1967). On the Kolmogorov-Smirnov test for normality

with mean and variance unknown. *J. Amer. Statist. Assoc.* 62 399-

402.

LILLIEFORS, H. W. (1969). On the Kolmogorov-Smirnov test for the expo-

nential distribution with mean unknown. *J. Amer. Statist. Assoc.*

64 387-389.

MANN, N. R., R. E. SCHAFER, and N. D. SINGPURWALLA (1974). *Methods for*

Statistical Analysis of Reliability and Life Data. John Wiley and

Sons, Inc., New York.

MANN, N. R., E. M. SCHEUER, and K. W. FERTIG (1973). A new goodness
of fit test for the two parameter Weibull or extreme value distri-
bution with unknown parameters. *Comm. Statist.* 2 383-400.

MARSHALL, A. W. and F. PROSCHAN (1965). Maximum likelihood estimation
for distributions with monotone failure rate. *Ann. Math. Statist.*
36 69-77.

SERFLING, R. J. and C. L. WOOD (1976). On null hypothesis limiting
distributions of Kolmogorov-Smirnov type statistics with estimated
location and scale parameters. Technical report, Florida State
University, Tallahassee, Florida.

SHORACK, G. (1972). Convergence of quantile and spaceing process with
applications. *Ann. Math. Statist.* 43 1400-1411.

SINGPURWALLA, N. D. (1975). Time series analysis and forecasting of
failure rate processes. *Reliability and Fault Tree Analysis* (R. E.
Barlow, J. Fussell, and N. D. Singpurwalla, eds.). Conference volume
published by SIAM, Philadelphia.

STEPHENS, M. A. (1977). Goodness-of-fit for the extreme value distribu-
tion. *Biometrika* 64 583-588.

THOMPSON, W. A., Jr. (1976). Fisherman's luck. *Biometrics* 32 265-
271.

THE GEORGE WASHINGTON UNIVERSITY

BENEATH THIS PLAQUE
IS BURIED
A VAULT FOR THE FUTURE
IN THE YEAR 2056

THE STORY OF ENGINEERING IN THIS YEAR OF THE PLACING OF THE VAULT AND
ENGINEERING HOPES FOR THE TOMORROWS AS WRITTEN IN THE RECORDS OF THE
FOLLOWING GOVERNMENTAL AND PROFESSIONAL ENGINEERING ORGANIZATIONS AND
THOSE OF THIS GEORGE WASHINGTON UNIVERSITY.

BOARD OF COMMISSIONERS DISTRICT OF COLUMBIA
UNITED STATES ATOMIC ENERGY COMMISSION
DEPARTMENT OF THE ARMY UNITED STATES OF AMERICA
DEPARTMENT OF THE NAVY UNITED STATES OF AMERICA
DEPARTMENT OF THE AIR FORCE UNITED STATES OF AMERICA
NATIONAL ADVISORY COMMITTEE FOR AERONAUTICS
NATIONAL BUREAU OF STANDARDS U S DEPARTMENT OF COMMERCE
AMERICAN SOCIETY OF CIVIL ENGINEERS
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS
THE SOCIETY OF AMERICAN MILITARY ENGINEERS
AMERICAN INSTITUTE OF MINING & METALLURGICAL ENGINEERS
DISTRICT OF COLUMBIA SOCIETY OF PROFESSIONAL ENGINEERS, INC
THE INSTITUTE OF RADIO ENGINEERS, INC
THE CHEMICAL ENGINEERS CLUB OF WASHINGTON
WASHINGTON SOCIETY OF ENGINEERS
FAULKNER KINGSBURY & STENHOUSE – ARCHITECTS
CHARLES H TOMPKINS COMPANY – BUILDERS
SOCIETY OF WOMEN ENGINEERS
NATIONAL ACADEMY OF SCIENCES. NATIONAL RESEARCH COUNCIL

THE PURPOSE OF THIS VAULT IS INSPIRED BY AND IS DEDICATED TO
CHARLES HOOK TOMPKINS, DOCTOR OF ENGINEERING
BECAUSE OF HIS ENGINEERING CONTRIBUTIONS TO THIS UNIVERSITY, TO HIS
COMMUNITY, TO HIS NATION, AND TO OTHER NATIONS.

BY THE GEORGE WASHINGTON UNIVERSITY

ROBERT V. FLEMING
CHAIRMAN OF THE BOARD OF TRUSTEES

CLOYD H. MARVIN
PRESIDENT

JUNE THE TWENTIETH
1955

To cope with the expanding technology, our society must
be assured of a continuing supply of rigorously trained
and educated engineers. The School of Engineering and
Applied Science is completely committed to this ob-
jective.